UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/825,423	04/03/2001	Patricia C. Weber	ID01152	2057

24265 7590 10/03/2006

SCHERING-PLOUGH CORPORATION
PATENT DEPARTMENT (K-6-1, 1990)
2000 GALLOPING HILL ROAD
KENILWORTH, NJ 07033-0530

EXAMINER	
STEADMAN, DAVID J	
ART UNIT	**PAPER NUMBER**
1656	

DATE MAILED: 10/03/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/825,423	WEBER ET AL.	
	Examiner	Art Unit	
	David J. Steadman	1656	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *17 July 2006*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-3,7-9,11,21 and 22* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☒ Claim(s) *1-3,7 and 8* is/are allowed.

6)☒ Claim(s) *9,11,21 and 22* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☒ Other: *Appendices A, B, C*.

DETAILED ACTION

Application Status

1. Claims 1-3, 7-9, 11, and 21-22 are pending in the application.

2. Applicant's amendment to the claims, filed on 17 July 2006, is acknowledged. This listing of the claims replaces all prior versions and listings of the claims.

3. Applicant's amendment to the specification, filed on 17 July 2006, is acknowledged. In view of this statement, sequence compliance appears to be perfected.

4. Applicant's arguments filed on 17 July 2006 in response to the Office action mailed on 7 March 2006 have been fully considered and are deemed to be persuasive to overcome at least one of the rejections and/or objections previously applied. Rejections and/or objections not reiterated from previous office actions are hereby withdrawn.

5. The text of those sections of Title 35, U.S. Code not included in the instant action can be found in a prior Office action.

Claim Rejections - 35 USC § 112, Second Paragraph

6. Claims 9 and 21-22 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claims 9 and 21 are confusing in the recitation of "[a]n isolated polypeptide defined by a variant...the variant consists of a single amino acid substitution..."

According to MPEP 2111.03, "[t]he transitional phrase 'consisting of' excludes any element, step, or ingredient not specified in the claim." Thus, claims 9 and 21 would appear to read on a polypeptide variant consisting of a single amino acid as defined by the claims. Claim 22 is also rejected as being confusing in the use of the transitional phrase "consists of" in the recitation of "[a]n isolated polypeptide defined by a variant" of SEQ ID NO:5 "wherein the variant consists of a substitution of the amino acids at positions 255-258." In the interest of advancing prosecution, claims 9 and 21 have been interpreted as meaning a variant of SEQ ID NO:3, 5, or 6 with a single amino acid mutation, wherein the mutation is at position 73 or 81. Claim 22 has been interpreted as meaning the polypeptide of SEQ ID NO:5, wherein amino acids 255-258 of SEQ ID NO:5 are replaced with SEQ ID NO:7, 8, 9, 10, 11, 12, 13, or 14. It is suggested that applicant clarify the meaning of the claims.

Claim Rejections - 35 USC § 112, First Paragraph

7. Claim 22 is rejected under 35 U.S.C. § 112, first paragraph, written description, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. This is a new matter rejection.

New claim 22 is drawn to a polypeptide "defined by a variant of...SEQ ID NO:5, wherein the variant consists of a substitution...at positions 255-258 with SEQ ID NO:7,8,9,10,11,12,13, or 14. MPEP § 2163 states, "when filing an amendment an

applicant should show support in the original disclosure for new or amended claims" and "[i]f the originally filed disclosure does not provide support for each claim limitation, or if an element which applicant describes as essential or critical is not claimed, a new or amended claim must be rejected under 35 U.S.C. 112, para. 1, as lacking adequate written description." Applicant points to original claim 7, p. 11 of the specification, and "by comparing the full-length sequence in SEQ ID NO:1 with the subdomain I,II fragment sequence in SEQ ID NO:5" (instant response at p. 5, bottom). The examiner has reviewed applicant's cited supporting disclosure and has aligned SEQ ID NO:1 against SEQ ID NO:5 (see Appendix C). However, this disclosure does not appear to support the claimed variant polypeptide. It is suggested that applicant show support for new claim 22. If applicant maintains that the cited disclosure supports claim 22 as written, applicant is requested to provide a detailed explanation as to how this cited disclosure supports the polypeptide of claim 22.

8. The written description rejection of claim 11 under 35 U.S.C. § 112, first paragraph, is maintained for the reasons of record and the reasons stated below. The rejection was fully explained in the prior Office action.

RESPONSE TO ARGUMENT: Applicant argues the rejection is overcome by claim amendment to "define the polypeptide in the crystalline composition by both a specific amino acid sequence and a specific set of structural coordinates."

Applicant's argument is not found persuasive. The examiner maintains the position that the specification fails to describe the genus of crystalline compositions of

claim 11. While the amendment to the claims limits the *polypeptide* of the composition as the recitation of "crystalline composition" in claim 11 does not specifically define any of the crystalline compositions that fall within its definition, particularly as the recitation of "crystalline composition" does not define any structural features commonly possessed by members of the genus of crystalline compositions of SEQ ID NO:17 that distinguish them from others. One skilled in the art therefore cannot, as one can do with a fully described genus, visualize or recognize the identity of the members of the genus of proteins in crystalline form. In this case, the structure(s) of the genus of *crystals* of the protein of SEQ ID NO:17 is completely undefined.

Applicant appears to take the position that by virtue of limiting the polypeptide of the crystalline composition to SEQ ID NO:17 having the structural coordinates of Table 5, the genus of crystals is adequately described, however, it is well-known in the art that a single polypeptide can crystallize into a plurality of distinct crystal forms, which one cannot predict *a priori* (see, *e.g.,* Aleshin et al. *FEBS Lett* 434:42-46, 1998). Thus, as noted in the prior Office action, the genus of crystals encompasses species that are widely variant, encompassing crystals of unliganded and liganded forms of SEQ ID NO:17, wherein the liganded form is in complex with *any* ligand(s). In this case, the specification discloses only a single representative species of the genus of recited crystalline compositions, *i.e.,* a protein crystal of SEQ ID NO:17 having space group $P2_1$ and unit cell dimensions a=34.8 Å, b=67.1 Å, c=58.4 Å, $\alpha=\gamma=90°$, and $\beta=101.3°$ (see particularly pp. 41-42 of the specification, which teaches crystallization of SEQ ID NO:17, which is amino acids 181-324 of the HCV NS3 helicase of SEQ ID NO:1), which

is undisputed by applicant. Other than these single species, the specification fails to describe any other crystals of SEQ ID NO:17 as encompassed by the claims. MPEP § 2163 states "[f]or inventions in an unpredictable art, adequate written description of a genus which embraces widely variant species cannot be achieved by disclosing only one species within the genus." As such, the single disclosed species of crystals of SEQ ID NO:17 fails to describe all crystals as encompassed by the claim.

Given the lack of description of a representative number of protein crystals, the specification fails to sufficiently describe the claimed invention in such full, clear, concise, and exact terms that a skilled artisan would recognize that applicant was in possession of the claimed invention.

9. The scope of enablement rejection of claim 11 under 35 U.S.C. 112, first paragraph, is maintained for the reasons of record and the reasons stated below. The rejection was fully explained in the prior Office action.

RESPONSE TO ARGUMENT: Applicant argues the rejection is overcome by claim amendment to "define the polypeptide in the crystalline composition by both a specific amino acid sequence and a specific set of structural coordinates."

Applicant's argument is not found persuasive. The examiner maintains the position that the specification fails to enable all crystals as broadly encompassed by the claim. While the examiner acknowledges the amendment to limit the polypeptide of the crystal to SEQ ID NO:17 having the structural coordinates of Table 5, claim 11 nonetheless broadly encompasses all crystals of SEQ ID NO:17, unliganded or

complexed with any ligand, having any space group, and any unit cell dimensions. The specification discloses only a single working example of the claimed crystal, *i.e.,* a crystal of SEQ ID NO:17 having space group $P2_1$ and unit cell dimensions a=34.8 Å, b=67.1 Å, c=58.4 Å, $\alpha=\gamma=90°$, and $\beta=101.3°$ (see particularly pp. 41-42 of the specification, which teaches crystallization of SEQ ID NO:17, which is amino acids 181-324 of the HCV NS3 helicase of SEQ ID NO:1). The specification fails to disclose any other working examples or guidance for making other protein crystals of SEQ ID NO:17 under any other conditions with an expectation of obtaining diffraction-quality crystals. As noted in the prior Office action – and undisputed by application – the state of the art at the time of the invention acknowledges a high level of unpredictability for making a protein crystal. For example, the reference of Branden et al. ("Introduction to Protein Structure Second Edition", Garland Publishing Inc., New York, 1999; cited in the prior Office action) teaches that "[c]rystallization is usually quite difficult to achieve" (p. 375) and that "[w]ell-ordered crystals...are difficult to grow because globular protein molecules are large, spherical, or ellipsoidal objects with irregular surfaces, and it is impossible to pack them into a crystal without forming large holes or channels between the individual molecules" (p. 374). Also, Drenth et al. ("Principles of X-ray Crystallography," Springer, New York, 1995; cited in the prior Office action) teaches that "[t]he science of protein crystallization is an underdeveloped area" and "[p]rotein crystallization is mainly a trial-and-error procedure" (p. 1). One cannot predict *a priori* those conditions that will lead to the successful crystallization of a diffraction-quality crystal nor can one predict the space group symmetry or unit cell dimensions of the

resulting crystal. As stated above, even a single polypeptide can have multiple crystal forms, however, what form will result from which particular crystallization conditions – if any – remains highly unpredictable as evidenced by the state of the art at the time of the invention. While applicant may argue that a crystal of SEQ ID NO:17 in complex with a ligand or ligands can be prepared according to the disclosed method and would have the same space group and unit cell dimensions, there is no way to predict *a priori* the space group and unit cell dimensions of a protein, as evidenced by the references of Kierzek et al. (cited in the prior Office action; see cited relevant teachings). While methods of protein crystallography were known at the time of the invention, it was not routine in the art to make all polypeptide crystals as encompassed by the claims and screen for those that are diffraction-quality under any crystallization conditions as encompassed by the claims, diffract those crystals, and to determine those polypeptide crystal structures that represent biologically-relevant macromolecules.

In view of the overly broad scope of the claims, the lack of guidance and working examples provided in the specification, the high level of unpredictability as evidenced by the prior art, and the amount of experimentation required to make and use all crystals and make and use all three-dimensional structures and methods of "rational drug design" as broadly encompassed by the claims, undue experimentation would be necessary for a skilled artisan to make and use the entire scope of the claimed invention.

Conclusion

10. Status of the claims:

Claims 1-3, 7-9, 11, and 21-22 are pending.

Claims 1-3 and 7-8 appear to be in a condition for allowance.

Claims 9, 11, and 21-22 are rejected.

Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to David J. Steadman whose telephone number is 571-272-0942. The examiner can normally be reached on Mon to Fri, 7:30 am to 4:00 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Kathleen Kerr can be reached on 571-272-0931. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

David J. Steadman, Ph.D.
Primary Examiner
Art Unit 1656

APPENDIX A

```
CLUSTAL W (1.83) Multiple Sequence Alignments
Sequence type explicitly set to Protein
Sequence format is Pearson
Sequence 1: SEQ_ID_NO_3       435 aa
Sequence 2: SEQ_ID_NO_5       845 aa
Sequence 3: SEQ_ID_NO_6       708 aaSequence 4:
SEQ_ID_NO_17      423 aa
Start of Pairwise alignments
Sequences (1:2) Aligned. Score: 100
Sequences (1:3) Aligned. Score: 100
Sequences (1:4) Aligned. Score: 100
Sequences (2:2) Aligned. Score: 100
Sequences (2:3) Aligned. Score: 69.9153
Sequences (2:4) Aligned. Score: 100
Sequences (3:2) Aligned. Score: 69.9153
Sequences (3:3) Aligned. Score: 100
Sequences (3:4) Aligned. Score: 100
Sequences (4:2) Aligned. Score: 100
Sequences (4:3) Aligned. Score: 100
Sequences (4:4) Aligned. Score: 100
Start of Multiple Alignment


SEQ_ID_NO_6       GLYSERHISMETSERPRVALPHETHRASPASNSERSERPRPRALAVALPRGLNSERPHEG
SEQ_ID_NO_17      ------------SERPRVALPHETHRASPASNSERSERPRPRALAVALPRGLNSERPHEG
SEQ_ID_NO_3       GLYSERHISMETSERPRVALPHETHRASPASNSERSERPRPRALAVALPRGLNSERPHEG
SEQ_ID_NO_5       GLYSERHISMETSERPRVALPHETHRASPASNSERSERPRPRALAVALPRGLNSERPHEG
                          **************************************************

SEQ_ID_NO_6       LNVALALAHISLEUHISALAPRTHRGLYSERGLYLYSSERTHRLYSVALPRALAALATYR
SEQ_ID_NO_17      LNVALALAHISLEUHISALAPRTHRGLYSERGLYLYSSERTHRLYSVALPRALAALATYR
SEQ_ID_NO_3       LNVALALAHISLEUHISALAPRTHRGLYSERGLYLYSSERTHRLYSVALPRALAALATYR
SEQ_ID_NO_5       LNVALALAHISLEUHISALAPRTHRGLYSERGLYLYSSERTHRLYSVALPRALAALATYR
                  ************************************************************

SEQ_ID_NO_6       ALAALAGLNGLYTYRLYSVALLEUVALLEUASNPRSERVALALAALATHRLEUGLYPHEG
SEQ_ID_NO_17      ALAALAGLNGLYTYRLYSVALLEUVALLEUASNPRSERVALALAALATHRLEUGLYPHEG
SEQ_ID_NO_3       ALAALAGLNGLYTYRLYSVALLEUVALLEUASNPRSERVALALAALATHRLEUGLYPHEG
SEQ_ID_NO_5       ALAALAGLNGLYTYRLYSVALLEUVALLEUASNPRSERVALALAALATHRLEUGLYPHEG
                  ************************************************************

SEQ_ID_NO_6       LYALATYRMETSERLYSALAHISGLYVALASPPRASNILEARGTHRGLYVALARGTHRIL
SEQ_ID_NO_17      LYALATYRMETSERLYSALAHISGLYVALASPPRASNILEARGTHRGLYVALARGTHRIL
SEQ_ID_NO_3       LYALATYRMETSERLYSALAHISGLYVALASPPRASNILEARGTHRGLYVALARGTHRIL
SEQ_ID_NO_5       LYALATYRMETSERLYSALAHISGLYVALASPPRASNILEARGTHRGLYVALARGTHRIL
                  ************************************************************

SEQ_ID_NO_6       ETHRTHRGLYSERPRILETHRTYRSERTHRTYRGLYLYSPHELEUALAASPGLYGLYCYS
SEQ_ID_NO_17      ETHRTHRGLYSERPRILETHRTYRSERTHRTYRGLYLYSPHELEUALAASPGLYGLYCYS
SEQ_ID_NO_3       ETHRTHRGLYSERPRILETHRTYRSERTHRTYRGLYLYSPHELEUALAASPGLYGLYCYS
SEQ_ID_NO_5       ETHRTHRGLYSERPRILETHRTYRSERTHRTYRGLYLYSPHELEUALAASPGLYGLYCYS
                  ************************************************************

SEQ_ID_NO_6       SERGLYGLYALATYRASPILEILEILECYSASPGLUCYSHISSERTHRASPALATHRSER
SEQ_ID_NO_17      SERGLYGLYALATYRASPILEILEILECYSASPGLUCYSHISSERTHRASPALATHRSER
SEQ_ID_NO_3       SERGLYGLYALATYRASPILEILEILECYSASPGLUCYSHISSERTHRASPALATHRSER
SEQ_ID_NO_5       SERGLYGLYALATYRASPILEILEILECYSASPGLUCYSHISSERTHRASPALATHRSER
                  ************************************************************

SEQ_ID_NO_6       ILELEUGLYILEGLYTHRVALLEUASPGLNALAGLUTHRALAGLYALAARGLEUVALVAL
SEQ_ID_NO_17      ILELEUGLYILEGLYTHRVALLEUASPGLNALAGLUTHRALAGLYALAARGLEUVALVAL
SEQ_ID_NO_3       ILELEUGLYILEGLYTHRVALLEUASPGLNALAGLUTHRALAGLYALAARGLEUVALVAL
SEQ_ID_NO_5       ILELEUGLYILEGLYTHRVALLEUASPGLNALAGLUTHRALAGLYALAARGLEUVALVAL
                  ************************************************************
```

```
SEQ_ID_NO_6       LEUALATHRALATHRPRPRGLYSERGLYMETPHEASPSERSERVALLEU-----------
SEQ_ID_NO_17      LEUALATHRALATHR----------------------------------------------
SEQ_ID_NO_3       LEUALATHRALATHR----------------------------------------------
SEQ_ID_NO_5       LEUALATHRALATHRPRPRGLYSERVALTHRVALPRHISPRASNILEGLUGLUVALALAL
                  ***************

SEQ_ID_NO_6       --------CYSGLU--CYSTYRASPALAGLYCYSALATRPTYRG----------------
SEQ_ID_NO_17      ------------------------------------------------------------
SEQ_ID_NO_3       ------------------------------------------------------------
SEQ_ID_NO_5       EUSERTHRTHRGLYGLUILEPRPHETYRGLYLYSALAILEPRLEUGLUVALILELYSGLY


SEQ_ID_NO_6       ---------LU----------------------------LEUTHRPRALAGLUTHRTHR
SEQ_ID_NO_17      ------------------------------------------------------------
SEQ_ID_NO_3       ------------------------------------------------------------
SEQ_ID_NO_5       GLYARGHISLEUILEPHECYSHISSERLYSLYSLYSCYSASPGLULEUALAALALYSLEU


SEQ_ID_NO_6       VALARGLEU------ARGALATYRMETASNTHRPRGLYLEU---------------PRV
SEQ_ID_NO_17      ------------------------------------------------------------
SEQ_ID_NO_3       ------------------------------------------------------------
SEQ_ID_NO_5       VALALALEUGLYILEASNALAVALALATYRTYRARGGLYLEUASPVALSERVALILEPRT


SEQ_ID_NO_6       ALCYSGLNASPHISLEU------GLUPHETRPGLU------GLYVALPHETHRGLYLEU-
SEQ_ID_NO_17      ------------------------------------------------------------
SEQ_ID_NO_3       ------------------------------------------------------------
SEQ_ID_NO_5       HRASNGLYASPVALVALVALVALALATHRASPALALEUMETTHRGLYPHETHRGLYASPP


SEQ_ID_NO_6       -----THRHISILEASPALAHISPHELEU----------------------SERGLNTH
SEQ_ID_NO_17      ------------------------------------------------------------
SEQ_ID_NO_3       ------------------------------------------------------------
SEQ_ID_NO_5       HEASPSERVALILEASPCYSASNTHRSERASPGLYLYSPRGLNASPALAVALSERARGTH


SEQ_ID_NO_6       RLYSGLNSERGLYGLUASNPHEPRTYRLEUVALALATYRGLNALATHRVALCYSALAARG
SEQ_ID_NO_17      ------------------------------------------------------------
SEQ_ID_NO_3       ------------------------------------------------------------
SEQ_ID_NO_5       RGLNARGARGGLYARGTHRGLYARGGLYLYSPRGLYILETYRARGPHEVALALAPRGLYG


SEQ_ID_NO_6       ALAGLN
SEQ_ID_NO_17      ------
SEQ_ID_NO_3       ------
SEQ_ID_NO_5       LUARG-
```

APPENDIX B

Seq1 is SEQ ID NO:1, Seq2 is SEQ ID NO:17

```
 s-w opt: 2614  Z-score: 3193.8  bits: 602.0 E(): 4.3e-176
Smith-Waterman score: 2614;  100.000% identity (100.000% ungapped) in 412 aa overlap (510-921:1-
412)

     480       490       500       510       520       530
Seq1   EILEPRVALGLASNLEGLTHRTHRMETARGSERPRVALPHETHRASPASNSERSERPRPR
                                     ::::::::::::::::::::::::::::::
Seq2                                   SERPRVALPHETHRASPASNSERSERPRPR
                                               10        20        30

     540       550       560       570       580       590
Seq1   ALAVALPRGLNSERPHEGLNVALALAHISLEHISALAPRTHRGLYSERGLYLYSSERTHR
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2   ALAVALPRGLNSERPHEGLNVALALAHISLEHISALAPRTHRGLYSERGLYLYSSERTHR
               40        50        60        70        80        90

     600       610       620       630       640       650
Seq1   LYSVALPRALAALATYRALAALAGLNGLYTYRLYSVALLEVALLEASNPRSERVALALAA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2   LYSVALPRALAALATYRALAALAGLNGLYTYRLYSVALLEVALLEASNPRSERVALALAA
             100       110       120       130       140       150

     660       670       680       690       700       710
Seq1   LATHRLEGLYPHEGLYALATYRMETSERLYSALAHISGLYVALASPPRASNILEARGTHR
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2   LATHRLEGLYPHEGLYALATYRMETSERLYSALAHISGLYVALASPPRASNILEARGTHR
             160       170       180       190       200       210

     720       730       740       750       760       770
Seq1   GLYVALARGTHRILETHRTHRGLYSERPRILETHRTYRSERTHRTYRGLYLYSPHELEAL
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2   GLYVALARGTHRILETHRTHRGLYSERPRILETHRTYRSERTHRTYRGLYLYSPHELEAL
             220       230       240       250       260       270

     780       790       800       810       820       830
Seq1   AASPGLYGLYCYSSERGLYGLYALATYRASPILEILEILECYSASPGLCYSHISSERTHR
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2   AASPGLYGLYCYSSERGLYGLYALATYRASPILEILEILECYSASPGLCYSHISSERTHR
             280       290       300       310       320       330

     840       850       860       870       880       890
Seq1   ASPALATHRSERILELEGLYILEGLYTHRVALLEASPGLNALAGLTHRALAGLYALAARG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2   ASPALATHRSERILELEGLYILEGLYTHRVALLEASPGLNALAGLTHRALAGLYALAARG
             340       350       360       370       380       390

     900       910       920       930       940       950
Seq1   LEVALVALLEALATHRALATHRPRPRGLYSERVALTHRVALPRHISPRASNILEGLGLVA
       :::::::::::::::::::::
Seq2   LEVALVALLEALATHRALATHR
             400       410
```

APPENDIX C

```
>>Seq2                                                    (820 aa)
 s-w opt: 4989  Z-score: 6113.3  bits: 1143.2 E():    0
Smith-Waterman score: 4989;  93.103% identity (98.901% ungapped) in 870 aa overlap (496-1364:1-
820)

        470       480       490       500       510       520
Seq1    LYSALAVALASPPHEILEPRVALGLASNLEGLTHRTH-RMETARGSERPRVALPHETHRA
                                  ::  . :  :::   ::::::::::::::
Seq2                              GLYSERHISMET---SERPRVALPHETHRA
                                         10          20

          530       540       550       560       570       580
Seq1    SPASNSERSERPRPRALAVALPRGLNSERPHEGLNVALALAHISLEHISALAPRTHRGLY
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2    SPASNSERSERPRPRALAVALPRGLNSERPHEGLNVALALAHISLEHISALAPRTHRGLY
         30        40        50        60        70        80

          590       600       610       620       630       640
Seq1    SERGLYLYSSERTHRLYSVALPRALAALATYRALAALAGLNGLYTYRLYSVALLEVALLE
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2    SERGLYLYSSERTHRLYSVALPRALAALATYRALAALAGLNGLYTYRLYSVALLEVALLE
         90        100       110       120       130       140

          650       660       670       680       690       700
Seq1    ASNPRSERVALALAALATHRLEGLYPHEGLYALATYRMETSERLYSALAHISGLYVALAS
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2    ASNPRSERVALALAALATHRLEGLYPHEGLYALATYRMETSERLYSALAHISGLYVALAS
        150       160       170       180       190       200

          710       720       730       740       750       760
Seq1    PPRASNILEARGTHRGLYVALARGTHRILETHRTHRGLYSERPRILETHRTYRSERTHRT
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2    PPRASNILEARGTHRGLYVALARGTHRILETHRTHRGLYSERPRILETHRTYRSERTHRT
        210       220       230       240       250       260

          770       780       790       800       810       820
Seq1    YRGLYLYSPHELEALAASPGLYGLYCYSSERGLYGLYALATYRASPILEILEILECYSAS
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2    YRGLYLYSPHELEALAASPGLYGLYCYSSERGLYGLYALATYRASPILEILEILECYSAS
        270       280       290       300       310       320

          830       840       850       860       870       880
Seq1    PGLCYSHISSERTHRASPALATHRSERILELEGLYILEGLYTHRVALLEASPGLNALAGL
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2    PGLCYSHISSERTHRASPALATHRSERILELEGLYILEGLYTHRVALLEASPGLNALAGL
        330       340       350       360       370       380

          890       900       910       920       930       940
Seq1    THRALAGLYALAARGLEVALVALLEALATHRALATHRPRPRGLYSERVALTHRVALPRHI
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2    THRALAGLYALAARGLEVALVALLEALATHRALATHRPRPRGLYSERVALTHRVALPRHI
        390       400       410       420       430       440

          950       960       970       980       990      1000
Seq1    SPRASNILEGLGLVALALALESERTHRTHRGLYGLILEPRPHETYRGLYLYSALAILEPR
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2    SPRASNILEGLGLVALALALESERTHRTHRGLYGLILEPRPHETYRGLYLYSALAILEPR
        450       460       470       480       490       500

         1010      1020      1030      1040      1050      1060
Seq1    LEGLVALILELYSGLYGLYARGHISLEILEPHECYSHISSERLYSLYSLYSCYSASPGLL
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2    LEGLVALILELYSGLYGLYARGHISLEILEPHECYSHISSERLYSLYSLYSCYSASPGLL
        510       520       530       540       550       560
```

```
          1070      1080      1090      1100      1110      1120
Seq1  EALAALALYSLEVALALALEGLYILEASNALAVALALATYRTYRARGGLYLEASPVALSE
      :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2  EALAALALYSLEVALALALEGLYILEASNALAVALALATYRTYRARGGLYLEASPVALSE
      570       580       590       600       610       620

          1130      1140      1150      1160      1170      1180
Seq1  RVALILEPRTHRASNGLYASPVALVALVALALATHRASPALALEMETTHRGLYPHET
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2  RVALILEPRTHRASNGLYASPVALVALVALALATHRASPALALEMETTHRGLYPHET
      630       640       650       660       670       680

          1190      1200      1210      1220      1230      1240
Seq1  HRGLYASPPHEASPSERVALILEASPCYSASNTHRCYSVALTHRGLNTHRVALASPPHES
      ::::::::::::::::::::::::::::::::::                      :
Seq2  HRGLYASPPHEASPSERVALILEASPCYSASNTHR----------------------S
      690       700       710       720

          1250      1260      1270      1280      1290      1300
Seq1  ERLEASPPRTHRPHETHRILEGLTHRTHRTHRLEPRGLNASPALAVALSERARGTHRGLN
      ::  :::              ::       .   :::::::::::::::::::::::::
Seq2  ER--ASP-------------GL-------YLYSPRGLNASPALAVALSERARGTHRGLN
                          730               740       750       760

          1310      1320      1330      1340      1350      1360
Seq1  ARGARGGLYARGTHRGLYARGGLYLYSPRGLYILETYRARGPHEVALALAPRGLYGLARG
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Seq2  ARGARGGLYARGTHRGLYARGGLYLYSPRGLYILETYRARGPHEVALALAPRGLYGLARG
            770       780       790       800       810       820

          1370      1380      1390      1400      1410      1420
Seq1  PRSERGLYMETPHEASPSERSERVALLECYSGLCYSTYRASPALAGLYCYSALATRPTYR
```